Large scale surface water change observed by Sentinel-2 during the 2018 drought in Germany

Marc Wieland [1] *, Sandro Martinis [1]

[1] *German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Oberpfaffenhofen, D-82234 Wessling, Germany*

** Correspondence: marc.wieland@dlr.de*

Large scale surface water change observed by Sentinel-2 during the 2018 drought in Germany

Monitoring and understanding the spatio-temporal dynamics of hydrological droughts with seamless geographical coverage over large areas is essential for an assessment of impacts on water resources, industry, transport and human health. This became particularly relevant during the heat and drought of 2018 in Germany, which affected the country across various sectors and caused significant interruptions to ship traffic on rivers and lakes with negative impacts on tourism, transportation and supply chains. In this study, we provide a spatially and temporally consistent view on the 2018 hydrological drought in Germany as seen from Sentinel-2 satellite images. We extract water bodies with national coverage at different timestamps using an automated processing chain, which is based on a convolutional neural network and has originally been developed for near-real time flood monitoring. The method produces water segmentations with consistently high Overall Accuracy (≥0.95) and Kappa Coefficient (≥0.89), despite varying topography, land-use / land-cover and atmospheric conditions. Furthermore, we identify hotspots of change in water extent at national scale by comparing monthly water maps for 2018 with the respective maps of the previous year 2017. For the change hotspots we map change gradients and produce water extent time-series with mapping frequencies <5 days along a timeline of 12 months.

Keywords: Water segmentation; drought; change detection; Germany; Sentinel-2

1 Introduction

Drought monitoring and reconstruction initiatives largely deal with meteorological drought, which is the lack of precipitation and possibly increased evapotranspiration, as well as agricultural drought as a consequence to this deficiency (Horion et al. 2012; Zink et al. 2016). Hydrological drought, which is defined as below normal water discharge, evolves over longer time-periods from meteorological drought and is commonly more heterogeneous in space and time due to close interactions with hydrological preconditions (Hanel et al. 2018). Monitoring and understanding the spatio-temporal dynamics of hydrological droughts with seamless geographical coverage over large areas is essential in order to assess impacts on water resources, industry, transport and human health. This became particularly relevant during the 2018 drought in Germany. The months from February to July 2018 were extremely dry and the nation-wide mean precipitation in this period reached only 61% of the usual amount of rain. Moreover, the months from April to July 2018 were the warmest ever recorded in Germany since 1881, with mean temperatures in all districts exceeding the long-time reference value 1981 to 2010 by 1.7 K (Mühr et al. 2018). The increasingly high temperatures led to 89% of the German territory being under severe drought by mid-August ("German Drought Monitor - Helmholtz-Centre for Environmental Research" 2018). These extreme weather conditions strongly affected the water discharge with consequences for agriculture, water resources, human health and ecosystem services (Mühr et al. 2018). Low water levels in particular restricted ship traffic on rivers and lakes with negative impacts on tourism, transportation and supply chains (Das Niedrigwasser 2018, 2019).

Conventional hydrological monitoring systems rely on the availability of dense

networks of rain and stream gauging stations, which provide measures of rainfall and water height at any given time. This makes them important tools for flood forecasting and warning (Klemas 2015) or river traffic management (Schweighofer 2014). However, data from gauging stations are only point-based observations that cannot fully represent the spatial distribution of surface water. In this regard remote sensing can provide complementary information on surface water extent over large geographical areas, at high temporal frequency and low cost. Satellite-based water monitoring can support water authorities with relevant information on hydrology in ungauged areas, support the development of infrastructure projects (e.g., retention areas, irrigation schemes) or help to understand the impacts of changes in hydrology on environment, economy and human health.

Surface water monitoring has long been an important topic in remote sensing and a detailed review of recent studies in that direction can be found in Huang et al. (2018). Existing automated methods to segment water bodies in satellite images can be categorized into rule-based systems and machine learning models. Simple rule-based approaches commonly threshold a water index, such as the Normalized Difference Water Index (NDWI) (Gao 1996), and the Modified Normalized Difference Water Index (MNDWI) (Xu 2006) or the Multi Band Water Index (MBWI) (Wang et al. 2018). Other popular water indices include the Automated Water Extraction Index (AWEI) (Feyisa et al. 2014), the High Resolution Water Index (HRWI) (Yao et al. 2015) or the WI_{2015} (Fisher, Flood, and Danaher 2016). More refined methods rely on complex rule-sets that use additional spectral or multi-temporal information and ancillary datasets (Pekel et al. 2016). Rule-based methods may be transparent and produce accurate results under specific conditions, but they largely lack generalization ability and transferability between sensors, geographies and scene properties. Machine learning methods learn water characteristics from a set of labelled samples (Hollstein et al. 2016). Recently, studies have proposed Convolutional Neural Networks (CNNs), which combine convolutional and pooling layers to learn features from images, for water segmentation in satellite images. These largely report superior accuracy and generalization ability of CNN, compared to rule-based and classical machine learning approaches with hand-crafted features (Isikdogan, Bovik, and Passalacqua 2017; Li, Martinis, and Wieland 2019; Yu et al. 2017).

Spatio-temporal monitoring of surface water dynamics is usually achieved by using multi-temporal change detection (Lu et al. 2004). A widely used change detection method in surface water monitoring is to compare categorical maps derived from satellite images acquired at different times to identify areas of change, such as water increase or decrease (El-Asmar and Hereher 2011; Du et al. 2012). The quality of this post-classification comparison method is directly related to the accuracy of the single categorical maps and therefore requires a highly accurate water segmentation method. Alternative change detection methods compare multi-temporal images directly by means of a change-vector analysis (Landmann et al. 2013; Huang et al. 2016) or use principal component analysis (Rokni et al. 2014). Ji et al. (2018) propose a global surface water change method based on Moderate Resolution Imaging Spectroradiometer (MODIS) daily reflectance time-series and ancillary datasets at 500 m resolution. Yao et al. (2019) constructed long-term time-series of Landsat imagery to map changes to a global database of lakes and reservoirs. Donchyts et al. (2016) mapped global surface water change over the past 30 years from Landsat imagery at 30 m resolution. Pekel et al. (2016) processed the whole Landsat data archive between 1984 and 2015 to produce the most detailed and comprehensive global surface water map to date. They also

derived products related to surface water changes, such as water occurrence, seasonality, change intensity or annual recurrence. Despite the availability of global long-term water monitoring efforts, there is still a need for rapid and flexible methods to compute changes at high frequency and for any given time range and location. Previous work of the authors has targeted this by providing a fully automated processing chain for flood monitoring (Wieland and Martinis 2019). The processing chain uses a globally trained CNN to segment water bodies in high-resolution multi-spectral satellite images with consistent quality across sensors, locations, seasons and events. Considered sensors include Landsat Thematic Mapper (TM), Landsat Enhanced Thematic Mapper (ETM+), Landsat Operational Land Imager (OLI) and Sentinel-2 Multispectral Instrument (MSI).

The objective of this study is to show the potentials of remote sensing to quickly derive relevant and consistent change products for large-scale surface water studies from freely available data. Specifically, we show the capabilities of our previously developed flood monitoring chain to quantify and monitor hydrological droughts over large areas and at high temporal frequency. We provide a spatially and temporally consistent view on the 2018 hydrological drought in Germany and identify hotspots of change in water extent at national scale by comparing water maps derived from Sentinel-2 satellite images for October 2018 with the same month in 2017. For selected hotspot locations we further analyse surface water changes with a mapping frequency of <5 days along a timeline of 12 months between the two reference dates. We provide independent accuracy assessments of the water maps.

2 Study area and data

We focus our study on inland surface water bodies across the territory of Germany that covers an area of approximately 357,580 km² (Figure 1). According to land-use / land-cover statistics, 51.6% of the territory is used for agriculture, 30.6% are covered by forests, 13.7% are occupied by settlements and transportation networks, whereas 2.4% are made up of water bodies (lakes, rivers and channels) and 1.7% of other land-use / land-cover (e.g., military compounds, mining areas, wetlands, etc.) (Umwelt und Landwirtschaft 2018, 2018). The five largest river basins in Germany (Rhine, Danube, Elbe, Weser and Ems) cover about 90% of Germany and show significant differences in climatology and hydrological regimes. The largest part of the country belongs to the cool / temperate climatic zone and is dominated by humid westerly winds. Geographically, the climate shows a gradient from a maritime influence in the northwest to a more continental climate in the south and east. Accordingly, the Rhine and Danube river basins are exposed to highest annual precipitation of over 1,000 mm, whereas the Elbe basin receives the lowest rainfall (Huang et al. 2013). Depending on altitude and climate of the basin, glacial, nival or pluvial flow regimes dominate. The tributaries of Danube and upper Rhine in the south receive large amounts of melting water from snow and glaciers, which results in highest water levels during the early summer months. In the central uplands the influence of rain and snow increases and combinations of nival and pluvial regimes dominate with highest water levels during the winter months. With decreasing altitude towards the north, pluvial regimes become more prominent and annual water flows are increasingly dominated by precipitation and evapotranspiration. The nation-wide mean annual precipitation for the reference timeframe 1981–2010 is 819 mm, with the highest rainfalls typically being recorded during the summer months between May and September. However, since the beginning of the 21st century, Europe has been exposed to a series of very hot and dry summers

(2003, 2010, 2013 and 2015) (Hanel et al. 2018), of which the recent heat and drought in 2018 showed most extensive spatial and temporal coverage in Germany with significant effects on the hydrology (Mühr et al. 2018; Das Niedrigwasser 2018, 2019).

To quantify changes in surface water extent during this recent drought event, we use Sentinel-2 images to consistently cover the territory of Germany in the reference months October 2017 and 2018. The images are acquired at processing level L1C, which are radiometrically calibrated, orthorectified and delivered in tiles of 100 × 100 km. 58 tiles for each reference month (116 tiles in total) are required for a full geographical coverage of Germany. For each month and image tile we pick the scene with the least cloud-cover and select neighbouring tiles as close as possible to each other in time. Additionally, for three hotspot locations we acquire all available Sentinel-2 images along the timeline of 12 months between the two reference dates (254 tiles in total).

Furthermore, we use a global long-term surface water occurrence (w) layer derived from Landsat imagery between 1984 and 2015 (Pekel et al. 2016). Water occurrence is computed as the sum of water detections per month divided by the sum of valid observations for the same month. From this dataset we extract the maximum water extent, which is defined as all locations ever being detected as water during 31 years ($w>0$), and use it as exclusion layer to focus our analysis on relevant water bodies. A buffer of 100 m is applied to make sure that we do not exclude any potentially relevant water pixels.

For accuracy assessment of the water segmentations we derive independent test datasets by applying a stratified random sampling with a minimum distance constraint of 2,000 m on the categorical outputs of the water segmentation algorithm for October 2017 and 2018 respectively. We use a fixed sample size of 500 per coverage and equally balance the samples across classes. Each point sample is manually interpreted by an expert in image analysis based on the respective Sentinel-2 image and labelled ‘Water’ or ‘Land’. Standard accuracy measures (Precision, Recall and F_1-score) are reported along with Overall Accuracy (OA) and Kappa Coefficient (κ) (Fawcett 2006). Precision is the ratio of true positive predictions to the total positive predictions. Recall is the ratio of true positive predictions to all observations in the actual class. F_1-score is the harmonic mean of Precision and Recall. OA is simply the ratio of correctly predicted observations to the total observations. A more robust metric is κ as it takes into account the possibility of the agreement occurring by chance.

3 Method

Figure 2 shows a schematic overview of the satellite image processing chain and its modules. Satellite images are downloaded from the Copernicus Open Access Hub (“Copernicus Open Access Hub” 2019) and fed into the process for preparation and analysis. Data are delivered as single band raster files with varying grid sizes. Therefore, we resample the lower resolution bands to match a grid size of 10 m. Six spectral bands are stacked together into a single dataset and transformed from Digital Numbers (DN) to Top of Atmosphere (TOA) reflectance by applying a sensor specific scale-factor. Specifically, we use bands Red, Green, Blue, Near-Infrared (NIR) and two Shortwave-Infrared bands (SWIR1 and SWIR2). Finally, for each reference month we create a seamless image mosaic for further analysis. The pre-processed image mosaics

are analysed with a globally trained convolutional neural network (CNN) that produces a semantic segmentation of the image into five classes ('Cloud', 'Cloud-Shadow', 'Water', 'Snow/Ice', 'Land'). The network is based on the U-Net architecture (Ronneberger, Fischer, and Brox 2015) and specifically considers cloud and cloud-shadow classes in order to identify pixels that may bias any downstream analysis and should therefore be marked 'Invalid'. A detailed description of the network architecture, reference dataset and training strategy is provided in Wieland and Martinis (2019).

We run the water segmentation on the image mosaics for October 2017 and 2018 separately. The network makes predictions on local windows with 256 × 256 pixels size, which may result in higher prediction errors towards the image borders. Therefore, during inference we expand the input image with mirror-padding, split it into overlapping tiles, run the predictions over batches of tiles, blend the prediction tiles to reconstruct the expanded input image's shape and un-pad the resulting prediction image. We use a tapered cosine window function to weight pixels when blending overlapping prediction tiles together. The final categorical output y is computed by maximizing the predicted probability vector $p(x)$. GPU inference speed is 0.41 seconds / megapixel, which means that a Sentinel-2 image tile at 10 m resolution with a typical size of 10,980 × 10,980 pixels can be analysed in less than one minute.

By reclassification of the categorical output we derive binary maps for water and valid pixels, where valid pixels are considered to belong to classes 'Water' or 'Land'. Invalid pixels are mapped equally throughout all timestamps, which means that if a pixel is identified as being invalid in one timestamp, it is also considered invalid in the other timestamp. The water maps are blended with the valid pixel map and the 31 years maximum water extent, which results in three classes ('Water', 'Land' and 'Invalid'). Finally, a water change image is derived by means of a post-classification comparison method (Lu et al. 2004). This results in four change classes: 'Constant' (water present in 2017 and 2018), 'Water loss' (water present in 2017 but not in 2018), 'Water gain' (water not present in 2017 but in 2018) and 'Invalid' (cloud or cloud-shadow in 2017 or 2018).

To aggregate the results for an identification of hotspots over large areas, we compute zonal statistics over a regular grid of hexagons with 10 km grid size that covers the whole study area. For each zone we compute the total water extent in 2017 and 2018 (in km²), the absolute water change between 2017 and 2018 (in km²) and the fraction of invalid to valid data (in %). We use hexagonal rather than rectangular grid cells, because they have a smaller perimeter-to-area ratio and hence reduce sampling bias related to edge effects (Birch, Sander, and Beecham 2007). Grid-cells that cover more than 50% invalid pixels, sea water or tidal zones are excluded.

For selected change hotspots we analyse all available satellite imagery between October 2017 and 2018, run a time-series analysis to show changes in water extent over time and map change gradients by means of relative water frequency. Relative water frequency is computed as sum of the binary water maps over the binary valid pixel maps, and provides an indication of the spatial dynamics of surface water within a given time range. A permanently water-covered location would be classified as belonging to class 'Water' in all valid observations throughout the time range, which would result in a relative water frequency f=1.0.

4 Results

Figure 3 shows water segmentation results for selected image tiles for October 2017 and 2018. Image tile 32UNE represents typical northern German lowlands around the Elbe estuary with the city of Hamburg in the centre. The area is dominated by maritime climate, whereas the image is taken during the dry period in October 2018 and shows large patches of bare soil, dry vegetation and low water levels. Image tile 32UPU focuses on south-eastern German alpine uplands, which are influenced by a continental climate and show transitions from hilly to mountainous topography. The scene depicts the area around Munich with major lakes Ammersee and Starnberger See and rivers Danube, Lech, and Isar. Image tile 32UMV shows a stretch of the Upper Rhine Rift between the cities of Karlsruhe and Mannheim in western Germany. The densely populated river flatlands are contrasted by central uplands in the east and west. Throughout all image tiles we can observe highly diverse land-use / land-cover, including different types of residential, industrial / commercial, agriculture, forest and bare soil classes. Water body types include rivers, canals, estuaries, lakes, ponds and harbours.

The predictions are able to delineate all major water bodies and even small objects, such as ponds or reservoirs, are detected and segmented with high precision. The qualitative visual evaluation of the results is confirmed by a quantitative accuracy assessment, for which the water maps have been tested against randomly selected and manually labelled point samples (Table 1). Water bodies are segmented with consistently high OA (≥0.95) and κ (≥0.89) despite the large geographical coverage with varying topography, land-use / land-cover and atmospheric conditions. Also differences in water conditions and extents between the two analysed timestamps do not significantly influence the performance. For a more in-depth evaluation of the generalization ability and global performance of the water and cloud / cloud-shadow segmentation algorithms the reader is referred to previous work of the authors (Wieland and Martinis 2019; Wieland, Li, and Martinis 2019).

Class	**October 2017**			**October 2018**		
	Precision	**Recall**	**F_1-score**	**Precision**	**Recall**	**F_1-score**
Water	0.99	0.91	0.95	0.98	0.95	0.97
Land	0.90	0.99	0.94	0.95	0.98	0.97
Total	**0.95**	**0.95**	**0.95**	**0.97**	**0.97**	**0.97**
OA	**0.95**			**0.97**		
κ	**0.89**			**0.94**		

Table 1: Accuracy assessment of the water maps for October 2017 and 2018 against independent test datasets with 500 samples per coverage.

Figure 4 and Figure 5 show water change between October 2017 and 2018 as change image at per-pixel level and as aggregated change zones at 10 km hexagonal grid cells. The water change zones allow to rapidly identifying change hotspots over large areas. At national scale one can observe major water declines along Rhine and Elbe rivers, and less severe water declines along Danube and Ems. Also the Lake of Constance and several standing water bodies in north-eastern Germany (e.g., Mecklenburger Seenplatte, Oberlausitz) were affected by water decline. Table 2 summarizes surface water changes at river basins on the territory of Germany. Across all river basins we can

observe a decrease of the area covered by inland surface water bodies between October 2017 and October 2018 by 344 km² (7.5%). The most severely affected river basins in terms of relative water decrease are Weser (-11.1%), Warnow /Peene (-10.1%) and Schlei / Trave (-9.4%). In terms of absolute water decrease the most affected river basins are Elbe (-110 km²), Rhine (-107 km²), Danube and Weser (each -34 km²).

River basin	October 2017 (km²)	October 2018 (km²)	Difference (km²)
Elbe	1670	1560	-110 (-6.6%)
Rhine	1161	1054	-107 (-9.2%)
Oder	213	198	-15 (-7.0%)
Ems	109	101	-8 (-7.3%)
Maas	36	34	-2 (-5.5%)
Warnow / Peene	227	204	-23 (-10.1%)
Eider	21	20	-1 (-4.7%)
Schlei / Trave	106	96	-10 (-9.4%)
Danube	738	704	-34 (-4.6%)
Weser	305	271	-34 (-11.1%)
Total	**4586**	**4242**	**-344 (-7.5%)**

Table 2: Summary of surface water changes on the territory of Germany between October 2017 and 2018 at river basin scale. Sea water, tidal zones and water bodies that were covered with clouds or cloud shadows in any of the two observation dates (invalid pixels) are not considered.

For further analysis, we select three hotspot locations with a water decline of more than 1.20 km² per grid cell. Namely, these are Elbe at Hohnstorf (A), Rhine at Koblenz (B) and the moor and pond landscape of Oberlausitz (C). Figure 6 shows the relative water frequencies for the three hotspots computed from all available Sentinel-2 images between October 2017 and 2018. The mean effective revisit period, which is the number of days between valid observations (under consideration of cloud coverage) (Wieland and Martinis 2019), shows small variations between the hotspots. For hotspot A we achieve a mean effective revisit period of 7.6 days (with a mean satellite revisit period of 4.5 days), for hotspot B 7.5 days (with a mean satellite revisit period of 4.7 days) and hotspot C 7.4 days (with a mean satellite revisit period of 4.3 days). For each acquisition date we compute the total water area detected within the hotspot and plot time-series of monthly mean water area. For a profile perpendicular to the main water flow across the water body at gauging station locations, we also plot time-series of monthly mean water extent.

Hotspot A (Elbe at Hohnstorf) shows the most dynamic surface water changes with larger patches of temporary flooded meadows along the main river. The maximum water extent within the observation period at gauging station Hohnstorf is reached in December 2017 (403 m), after which the water decreases almost linearly until the minimum extent in August 2018 (254 m). The time-series of water area for the whole hotspot shows a similar pattern with the largest water area being detected in January 2018 (41 km²) and the smallest area in August 2018 (13 km²). Time-series analysis of hotspot B (Rhine at Koblenz) reveals a smaller spatial extent of temporary flooded areas, but shows that relatively large branches of the main river dry out during the observation period. The maximum water extent at Koblenz station is reached in December 2017 (420 m), followed by a rapid water decrease until the minimum in March 2018 (300 m). The water extent slightly recovers in May / June, but reaches a

second minimum in August 2018 (325 m). The water area time-series for the whole hotspot also has its main peak in December 2017 (27 km²) and reaches its minimum in March 2018 (19 km²). Hotspot C (moor and pond landscape of Oberlausitz) largely covers standing water bodies and shows significant changes in water extent throughout its ponds and lakes. The profile for extent time-series analysis is located at the dam Quitzdorf, which blocks the river "Schwarze Schöps" and acts as service water reservoir and flood protection measure. Water extent at this location is regulated and shows an almost linear decrease throughout the observation period. The maximum water extent is reached in January 2018 (478 m) and the minimum extent in October 2018 (146 m). The area of all water bodies covered by this hotspot also decreases from October 2017 (47 km²) until June 2018 (35 km²) but shows a slight recovery after the minimum in June.

5 Discussion

In this study, we could show that Sentinel-2 data provided sufficiently high spatial and temporal resolutions to monitor surface water changes with high precision in short time, over large areas and free of (data) costs. We successfully extended an automated processing chain, which has originally been developed for flood monitoring, to quantify hydrological droughts. Compared to our previous work on flood monitoring, we have added a change detection method to outline water changes at per-pixel level and at aggregated grid zones. The proposed post-classification comparison method is simple but efficient and due to high accuracies of the input water maps it is also capable to produce accurate change results. Especially hydrological drought events affect surface water body extents at small scales with large spatio-temporal variability and coverage. We showed that it is possible to quantify such heterogeneous water dynamics and to identify drought hotspots with a bi-temporal change detection method. High frequency mapping of surface water extent for hotspots revealed more fine-grained surface water dynamics.

With Sentinel-2 we could achieve a mapping frequency of <5 days when considering the mean satellite revisit period over the study area. Since Sentinel-2 is an optical sensor and thus affected by cloud coverage, we also computed the mean effective revisit period, which specifically considers cloud and cloud shadow coverage. It is therefore a more meaningful measure to understand the applicability of an optical satellite sensor to provide valid observations during monitoring tasks. Since commonly not all pixels in all acquisitions are cloud and cloud shadow free, the effective revisit period of optical sensors is longer than the satellite revisit period. For the hotspot areas the mean effective revisit period between 2017-10-01 and 2018-10-31 was 7.5 days. It should be noted, however, that it is only valid for a given time and area of interest.

Test samples for the accuracy assessment of water segmentations are derived by manual interpretation of the respective satellite images. The results presented in this study are therefore to be regarded as relative to the performance of a human analyst and not to an absolute ground-truth. Results are, moreover, based on point samples, which may result in potentially higher accuracy values than a comparison against a complete manual segmentation at the pixel level. This positive bias is related to the fact that segmentations need to define a hard class boundary on often fuzzy outlines such as the land-water-border. Nevertheless, the presented accuracy values are well in line with previous accuracy assessments of the water segmentation method that were based on

pixel-by-pixel comparisons with globally sampled and manually segmented test areas (Wieland and Martinis 2019).

Despite an attempt to create cloudless image mosaics by carefully choosing the most appropriate images within pre-defined reference months, completely cloudless coverages for the whole study area could not be achieved. A more refined approach that blends image tiles at different acquisition dates over cloudy areas would be desirable to reduce or even completely remove cloud and cloud-shadow pixels. This would not only allow unbiased down-stream analysis, but would enable reporting of absolute statistics about water extent and change seamlessly over the whole study area. In the present study, the reported absolute statistics at larger aggregation scales of river basins or country are not comparable with national statistics, because we omit sea water, tidal zones and areas that are covered by clouds or cloud shadows in any of the two observation dates (invalid pixels). Nevertheless, relative changes can be considered representative and relevant for drought monitoring. Combining Sentinel-2 with other optical (e.g., Landsat OLI) or Synthetic Aperture Radar (SAR) data (e.g., Sentinel-1) could further aid the generation of cloud-free coverages at specific dates. It could, moreover, improve the effective revisit period and thus would allow for even finer grained temporal granularity of change hotspot analyses.

6 Conclusions

In this study, we applied a processing chain that has originally been developed for automated flood monitoring to systematically assess hydrological drought from Sentinel-2 images. The solution provided consistently accurate water segmentations ($\geq$0.95 OA and $\geq$0.89 κ) at high spatial resolution (10 m), over large geographical areas (>350,000 km²), at different times and despite large variations in environmental and atmospheric conditions. By adding a change detection module that compares the results of single water segmentations at per-pixel level and at aggregated hexagonal grid-zones, we were able to detect and quantify small scale changes in surface water extent as well as change hotspots at national scale. Time-series analysis of hotspot locations revealed further insights into the spatio-temporal dynamics of surface water changes.

The presented solution allowed to process raw satellite data into actionable information products that can provide rapid situational awareness in drought situations. We could show that remote sensing can provide complementary information on surface water extent over large geographical areas, at high temporal frequency and low costs. Satellite-based water monitoring can not only provide relevant information on hydrology in ungauged areas, but can also provide additional information about water extent in gauged areas. With their seamless coverage and systematic acquisition schedule, Sentinel-2 data allowed producing an unprecedented view on the hydrological drought 2018 in Germany. This may support water management authorities to better understand the impacts of changes in hydrology on environment, economy and human health.

Future work will focus on automating and improving the creation of cloudless image mosaics and to move towards continuous national and potentially continental monitoring of water bodies at high spatial and temporal resolution. Systematically integrating in-situ with satellite-based information will be of particular research interest in this context. This could be useful for early detection and surveillance of natural

disasters, such as droughts or floods, and could provide further insights into their causes, evolution and impacts over large areas.

7 References

Birch, Colin, Oom Sander, and Jonathan Beecham. 2007. "Rectangular and Hexagonal Grids Used for Observation, Experiment and Simulation in Ecology." *Ecological Modelling* 206 (3–4): 347–359.

"Copernicus Open Access Hub." 2019. https://scihub.copernicus.eu/.

Das Niedrigwasser 2018. 2019. Koblenz: Bundesanstalt für Gewässerkunde. http://doi.bafg.de/BfG/2019/Niedrigwasser_2018.pdf.

Donchyts, Gennadii, Fedor Baart, Hessel Winsemius, Noel Gorelick, Jaap Kwadijk, and Nick van de Giesen. 2016. "Earth's Surface Water Change over the Past 30 Years |." *Nature Climate Change* 6 (August): 810–813.

Du, Zhiqiang, Bin Linghu, Feng Ling, Wenbo Li, Weidong Tian, Hailei Wang, Yuanmiao Gui, Bingyu Sun, and Xiaoming Zhang. 2012. "Estimating Surface Water Area Changes Using Time-Series Landsat Data in the Qingjiang River Basin, China." *Journal of Applied Remote Sensing* 6 (1): 063609. doi:10.1117/1.JRS.6.063609.

El-Asmar, H. M., and M. E. Hereher. 2011. "Change Detection of the Coastal Zone East of the Nile Delta Using Remote Sensing." *Environmental Earth Sciences* 62 (4): 769–777. doi:10.1007/s12665-010-0564-9.

Fawcett, Tom. 2006. "An Introduction to ROC Analysis." *Pattern Recognition Letters* 27 (8): 861–874. doi:10.1016/j.patrec.2005.10.010.

Feyisa, Gudina L., Henrik Meilby, Rasmus Fensholt, and Simon R. Proud. 2014. "Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery." *Remote Sensing of Environment* 140 (January): 23–35. doi:10.1016/j.rse.2013.08.029.

Fisher, Adrian, Neil Flood, and Tim Danaher. 2016. "Comparing Landsat Water Index Methods for Automated Water Classification in Eastern Australia." *Remote Sensing of Environment* 175 (March): 167–182. doi:10.1016/j.rse.2015.12.055.

Gao, Bo-cai. 1996. "NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space." *Remote Sensing of Environment* 58 (3): 257–266. doi:10.1016/S0034-4257(96)00067-3.

"German Drought Monitor - Helmholtz-Centre for Environmental Research." 2018. https://www.ufz.de/duerremonitor/.

Hanel, Martin, Oldřich Rakovec, Yannis Markonis, Petr Máca, Luis Samaniego, Jan Kyselý, and Rohini Kumar. 2018. "Revisiting the Recent European Droughts from a Long-Term Perspective." *Scientific Reports* 8 (1): 9499. doi:10.1038/s41598-018-27464-4.

Hollstein, André, Karl Segl, Luis Guanter, Maximilian Brell, and Marta Enesco. 2016. "Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images." *Remote Sensing* 8 (8): 666. doi:10.3390/rs8080666.

Horion, S, H Carrão, A Singleton, P Barbosa, J Vogt, European Commission, Joint Research Centre, and Institute for Environment and Sustainability. 2012. *JRC Experience on the Development of Drought Information Systems: Europe, Africa and Latin America.* Luxembourg: OPOCE. http://dx.publications.europa.eu/10.2788/15761.

Huang, Chang, Yun Chen, Shiqiang Zhang, and Jianping Wu. 2018. “Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review.” *Reviews of Geophysics* 56 (2): 333–360. doi:10.1029/2018RG000598.

Huang, Chang, Xiaoyu Zan, Xuewen Yang, and Shiqiang Zhang. 2016. “Surface Water Change Detection Using Change Vector Analysis.” In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2834–2837. Beijing, China: IEEE. doi:10.1109/IGARSS.2016.7729732.

Huang, Shaochun, Fred F. Hattermann, Valentina Krysanova, and Axel Bronstert. 2013. “Projections of Climate Change Impacts on River Flood Conditions in Germany by Combining Three Different RCMs with a Regional Eco-Hydrological Model.” *Climatic Change* 116 (3–4): 631–663. doi:10.1007/s10584-012-0586-2.

Isikdogan, Furkan, Alan C. Bovik, and Paola Passalacqua. 2017. “Surface Water Mapping by Deep Learning.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10 (11): 4909–4918. doi:10.1109/JSTARS.2017.2735443.

Ji, Luyan, Peng Gong, Jie Wang, Jiancheng Shi, and Zhiliang Zhu. 2018. “Construction of the 500-m Resolution Daily Global Surface Water Change Database (2001–2016).” *Water Resources Research* 54 (12). doi:10.1029/2018WR023060.

Klemas, Victor. 2015. “Remote Sensing of Floods and Flood-Prone Areas: An Overview.” *Journal of Coastal Research* 314 (July): 1005–1013. doi:10.2112/JCOASTRES-D-14-00160.1.

Landmann, Tobias, Matthias Schramm, Christian Hüttich, and Stefan W. Dech. 2013. “MODIS-Based Change Vector Analysis for Assessing Wetland Dynamics in Southern Africa.” In . doi:10.1080/2150704X.2012.699201.

Li, Yu, Sandro Martinis, and Marc Wieland. 2019. “Urban Flood Mapping with an Active Self-Learning Convolutional Neural Network Based on TerraSAR-X Intensity and Interferometric Coherence.” *ISPRS Journal of Photogrammetry and Remote Sensing* 152 (June): 178–191. doi:10.1016/j.isprsjprs.2019.04.014.

Lu, D., P. Mausel, E. Brondízio, and E. Moran. 2004. “Change Detection Techniques.” *International Journal of Remote Sensing* 25 (12): 2365–2401. doi:10.1080/0143116031000139863.

Mühr, Bernhard, Susanne Kubisch, Andreas Marx, Johanna Stötzer, Christina Wisotzky, Christian Latt, Fabian Siegmann, Maren Glattfelder, Susanna Mohr, and Michael Kunz. 2018. “Dürre & Hitzewelle Sommer 2018 (Deutschland),” 19.

Pekel, Jean-François, Andrew Cottam, Noel Gorelick, and Alan S. Belward. 2016. “High-Resolution Mapping of Global Surface Water and Its Long-Term Changes.” *Nature* 540 (7633): 418–422. doi:10.1038/nature20584.

Rokni, Komeil, Anuar Ahmad, Ali Selamat, and Sharifeh Hazini. 2014. “Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery.” *Remote Sensing* 6 (5): 4173–4189. doi:10.3390/rs6054173.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. 2015. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, edited by Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, 9351:234–241. Cham: Springer International Publishing. doi:10.1007/978-3-319-24574-4_28.

Schweighofer, Juha. 2014. “The Impact of Extreme Weather and Climate Change on Inland Waterway Transport.” *Natural Hazards* 72 (1): 23–40. doi:10.1007/s11069-012-0541-6.

Umwelt und Landwirtschaft 2018. 2018. Dessau-Roßlau: Umweltbundesamt.

Wang, Xiaobiao, Shunping Xie, Xueliang Zhang, Cheng Chen, Hao Guo, Jinkang Du, and Zheng Duan. 2018. “A Robust Multi-Band Water Index (MBWI) for Automated Extraction of Surface Water from Landsat 8 OLI Imagery.” *International Journal of Applied Earth Observation and Geoinformation* 68 (June): 73–91. doi:10.1016/j.jag.2018.01.018.

Wieland, Marc, Yu Li, and Sandro Martinis. 2019. “Multi-Sensor Cloud and Cloud Shadow Segmentation with a Convolutional Neural Network.” *Remote Sensing of Environment* 230 (111203): 1–12.

Wieland, Marc, and Sandro Martinis. 2019. “A Modular Processing Chain for Automated Flood Monitoring from Multi-Spectral Satellite Data.” *Remote Sensing* 11 (9): 2330.

Xu, Hanqiu. 2006. “Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery.” *International Journal of Remote Sensing* 27 (14): 3025–3033. doi:10.1080/01431160600589179.

Yao, Fangfang, Chao Wang, Di Dong, Jiancheng Luo, Zhanfeng Shen, and Kehan Yang. 2015. “High-Resolution Mapping of Urban Surface Water Using ZY-3 Multi-Spectral Imagery.” *Remote Sensing* 7 (9): 12336–12355. doi:10.3390/rs70912336.

Yao, Fangfang, Jida Wang, Chao Wang, and Jean-François Crétaux. 2019. “Constructing Long-Term High-Frequency Time Series of Global Lake and Reservoir Areas Using Landsat Imagery.” *Remote Sensing of Environment* 232 (October): 111210. doi:10.1016/j.rse.2019.111210.

Yu, Long, Zhiyin Wang, Shengwei Tian, Feiyue Ye, Jianli Ding, and Jun Kong. 2017. “Convolutional Neural Networks for Water Body Extraction from Landsat Imagery.” *International Journal of Computational Intelligence and Applications* 16 (01): 1750001. doi:10.1142/S1469026817500018.

Zink, Matthias, Luis Samaniego, Rohini Kumar, Stephan Thober, Juliane Mai, David Schäfer, and Andreas Marx. 2016. “The German Drought Monitor.” *Environmental Research Letters* 11 (7): 074002. doi:10.1088/1748-9326/11/7/074002.

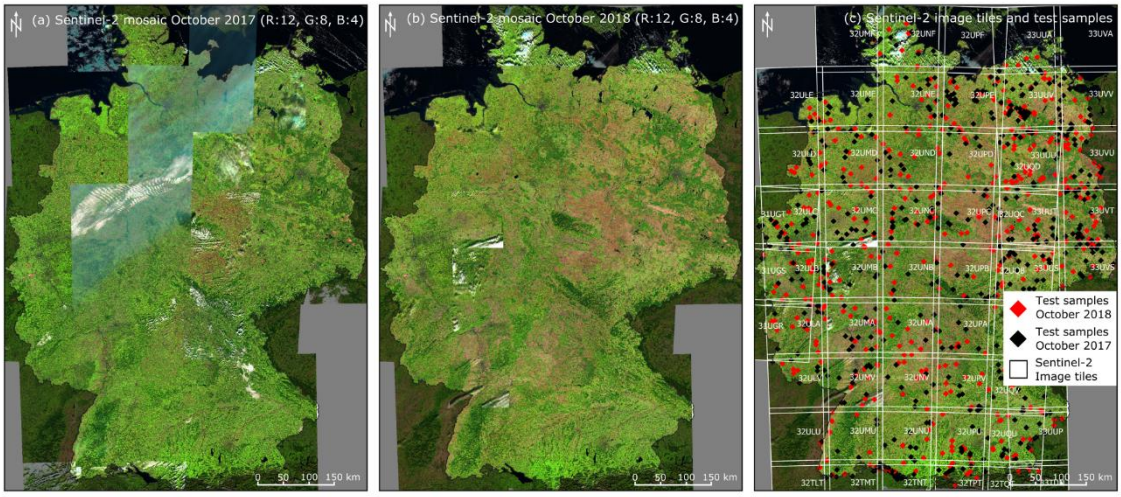

Figure 1: Overview of the study area with Sentinel-2 image mosaics for October 2017 and 2018. Image tile footprints superimposed on the 2018 image mosaic with locations of independent test samples for accuracy assessment.

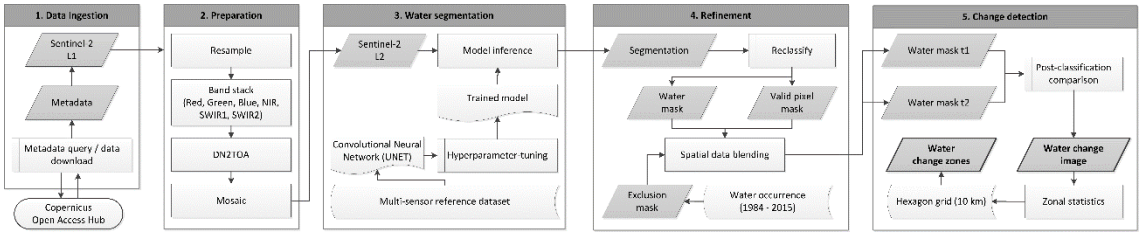

Figure 2: Scheme of the satellite image processing and analysis chain.

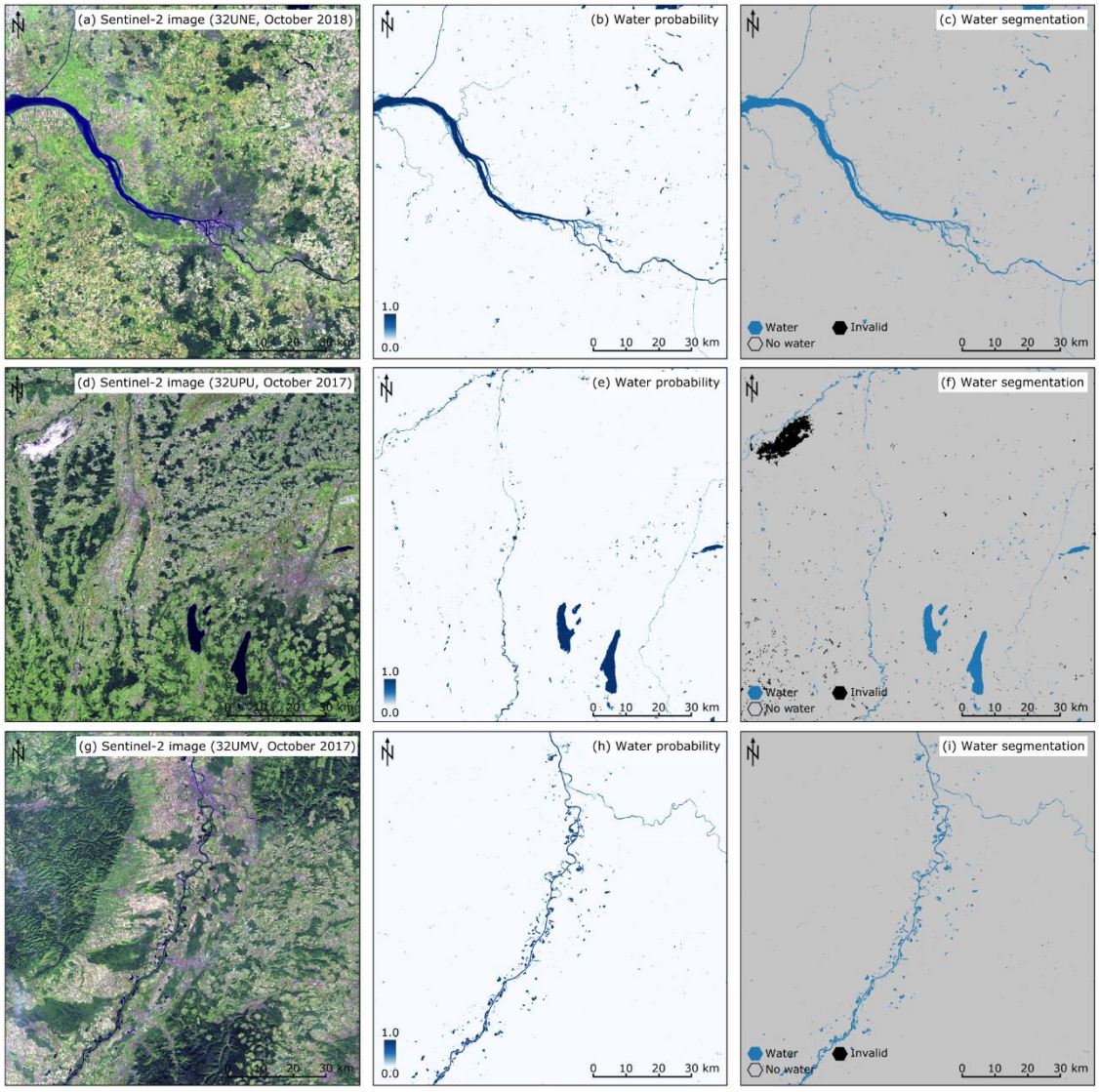

Figure 3: Selected Sentinel-2 image tiles, predicted water probabilities from the CNN and final water segmentations.

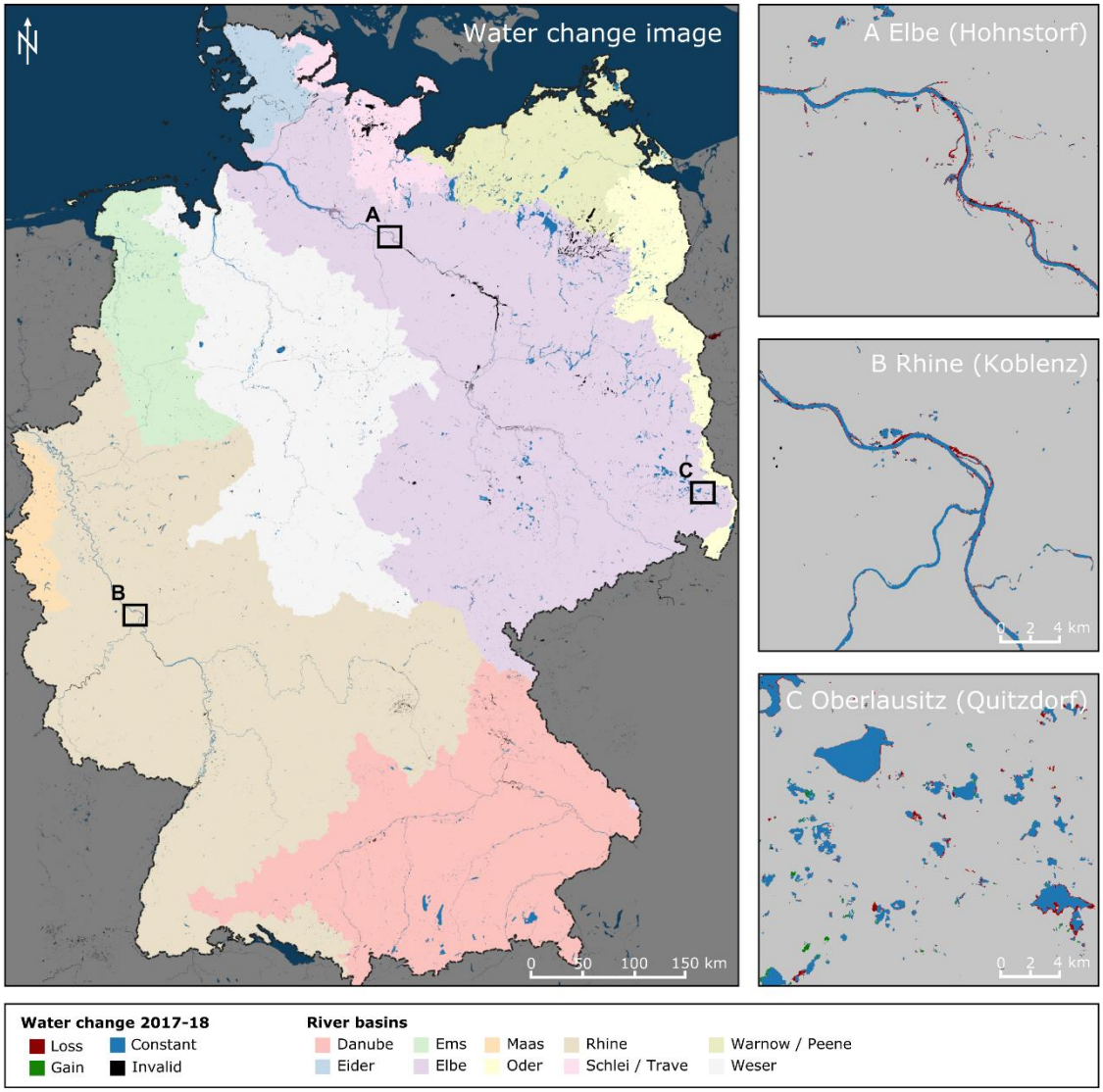

Figure 4: Per-pixel surface water change between October 2017 and 2018 with details for three change hotspots. The results are superimposed on major river basins.

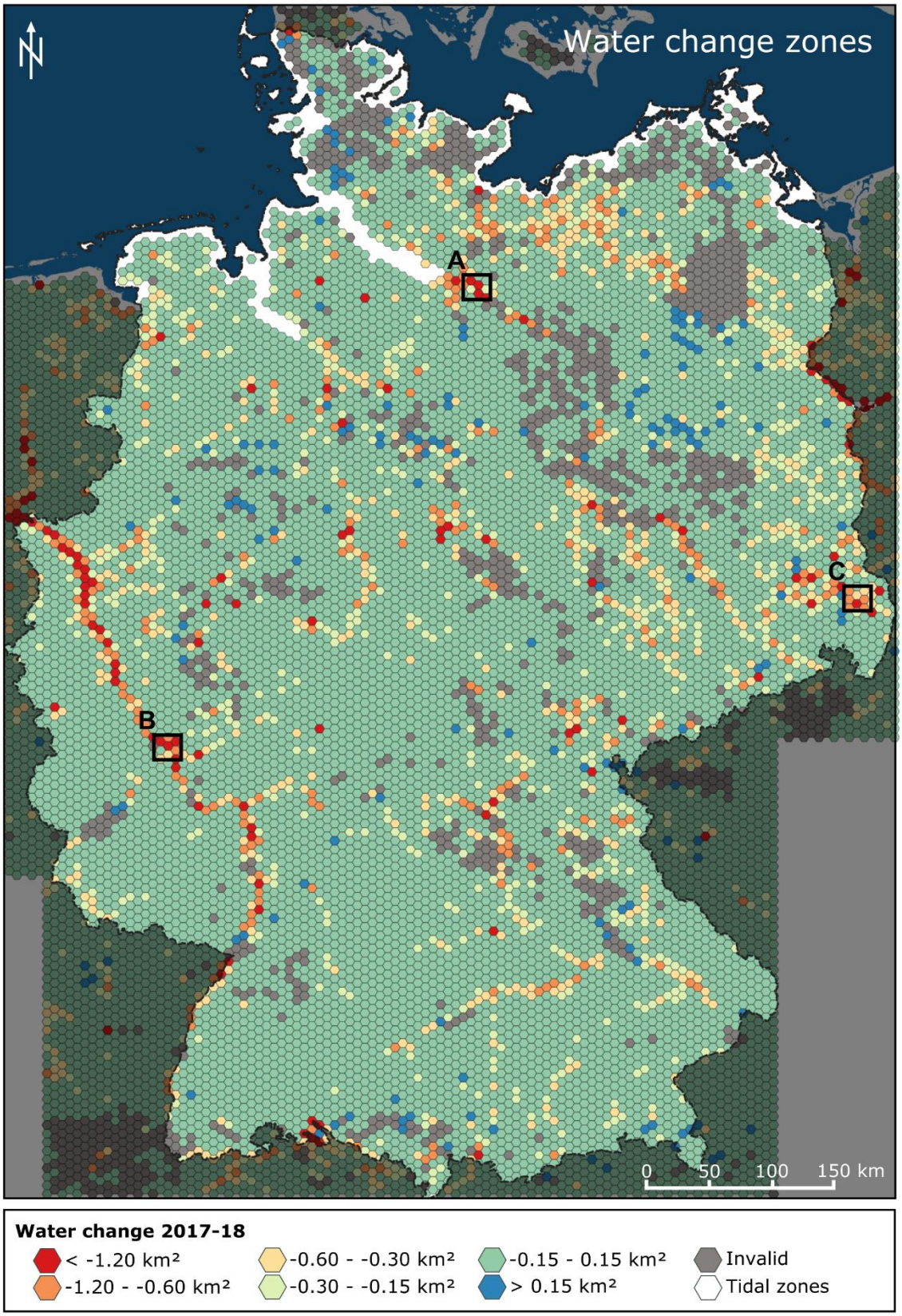

Figure 5: Aggregated surface water change between October 2017 and 2018 at hexagonal 10 km grid zones.

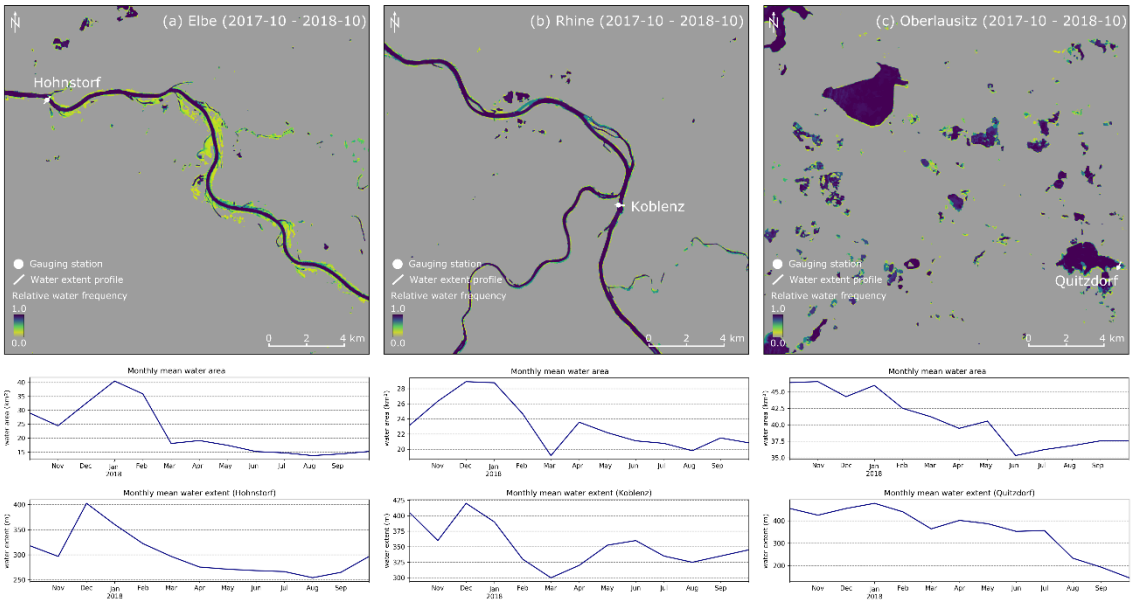

Figure 6: Relative water frequencies for three change hotspots. Timelines of monthly mean water extent at gauging station locations and timelines of monthly mean water area coverage per hotspot.